February 15, 2001

DSSD Census 2000 Procedures and Operations Memorandum Series U#-6, Revised

MEMORANDUM FOR	Ruth Ann Killion Chief, Planning, Research, and Evaluation Division
From:	Howard Hogan Chief, Decennial Statistical Studies Division
Subject:	Study Plan for the Analysis of Housing Unit Coverage in Puerto Rico, O.8.

Attached is the study plan for the Analysis of Housing Unit Coverage in Puerto Rico, O.8.
A new section of cost information has been added since the last time this study plan was issued.
The Census 2000 Evaluation Program quality assurance process was applied to the methodology
development and the study plan review process. The study plan is sound and appropriate for
completeness and accuracy, and it answers its intended category questions as appropriate.

If you have questions about this study plan, please contact Xijian Jim Liu on (301) 457-8325.

Attachment (Analysis of Housing Unit Coverage in Puerto Rico, O.8. Study Plan)

cc: Evaluations Executive Steering Committee
       Keith Bennett (PRED)
       Linda Brudvig
       Jason Machowski
       Rita Petroni
       Danny Childers (DSSD)

<div align="center">

Census 2000 Operational Summary
Study Plan
(Final)

</div>

I.   Name of Operation

Analysis of Housing Unit Coverage in Puerto Rico.
DSSD Census 2000 Evaluation Project O.8.

II.   Project Managers

Xijian Jim Liu (DSSD)
Telephone Number: (301) 457-8325
Email Address: xijian_liu@census.gov

III.   Operational Background

Census 2000

The Accuracy and Coverage Evaluation (A.C.E.) measures the overall and differential coverage of the US and Puerto Rico population in Census 2000. It uses dual system estimation which requires two samples, the P-sample and the E-sample. The P-sample consists of housing units enumerated independent of the census. The E-sample consists of housing units enumerated in the census. The major steps of the A.C.E. are housing unit matching and person matching. Housing units within the sample clusters were listed and matched to the January 2000 version of the Decennial Master Address File (DMAF). After matching and reconciliation, the P-sample yields an estimate of housing units missed in the census while the E-sample yields an estimate of housing units correctly enumerated in the census. Putting these two components together yields an A.C.E. estimate of census housing unit coverage.

In Puerto Rico, there was a separate A.C.E. listing sample that contains 559 block clusters and roughly 50,000 housing units. In the census, the entire population of Puerto Rico was enumerated using the update/leave method. The clusters in Puerto Rico were not eligible for medium and large block cluster reduction. After small block cluster subsampling, there were 499 clusters in Puerto Rico, including 36 small block clusters. Because there was additional information in the address fields in Puerto Rico that prevented the effective use of computer matching algorithm, housing units in Puerto Rico were not computer matched, but were matched clerically. The dual system estimation also has a separate set of post-strata for Puerto Rico.

<div align="center">

1

</div>

This study plan does not contain a literature review because there is no past research related to housing unit coverage in Puerto Rico.

IV.   Questions to be Answered and Methodology

The objective of this study is to measure the effectiveness of the DMAF in accurately reflecting all housing units existing in Puerto Rico on April 1, 2000.

1.   What is the net undercount rate of housing unit in Puerto Rico? How does the net undercount rate of housing units in Puerto Rico compare to that in the 50 states and the District of Columbia?

a.   Methodology

We will use dual-system estimation (*DSE*) to estimate the number of housing units.

$$DSE = \frac{C \cdot \frac{CE}{N_e}}{\frac{M}{N_p}}$$

where
C = the count of housing units in the census,
CE = the weighted estimate of the number of housing units correctly enumerated in the E-sample,
N_e = the weighted number of E-sample housing units,
M = the weighted number of P-sample housing unit matches,
N_p = the weighted number of P-sample housing units.

The dual system estimate of the number of housing units is computed within each post-stratum. The dual system estimate for the number of housing units in the entire Puerto Rico is the sum of estimated number of housing units in all post-strata.

The net undercount rate is $100\% \cdot (1 - \frac{C}{DSE})$.

The net undercount rate in Puerto Rico will be computed by selected variables and analyzed. The results of Puerto Rico will be compared with those of the 50 states and the District of Columbia provided by related project (O.3).

We will consider the following variables in this analysis:
tenure, MSA, single/multi-unit, return rates and other post-strata variables. (The variable MSA will have three values: San Juan CMSA, Other MSA,

2

and Non MSA. The list of post-strata variables for housing unit DSE will be finalized by the housing unit DSE team).

Standard errors will be computed using the stratified Jackknife method via VPLX.

    b.   Limitation

        See V.

    c.   Processing Requirements

        Programming and Computer

        Data files used for housing unit dual system estimation will be used.

        These data will be provided by DSSD and are expected to be available in June 2001.

2.   What is the non-match rate and erroneous enumeration rate of housing units in Puerto Rico? How do these rates in Puerto Rico compare to those in the 50 states and the District of Columbia?

    a.   Methodology

        The housing unit non-match rate is the weighted number of P-sample non-matched housing units divided by the weighted total of P-sample housing units.

        The housing unit erroneous enumeration rate is the weighted number of erroneously enumerated E-sample housing units divided by the weighted total of E-sample housing units.

        Housing unit non-match rate, and housing unit erroneous enumeration rate in Puerto Rico will be computed by selected variables and analyzed. The results of Puerto Rico will be compared with those of the 50 states and the District of Columbia provided by related projects (O.3 and O.10).

        We will consider the following variables in this analysis:
MSA, tenure, return rates and other post-strata variables.

        Standard errors will be computed using the stratified Jackknife method via VPLX.

b.	Limitation

See V.

c.	Processing Requirements

Programming and Computer

Data files used for housing unit dual system estimation will be used. ·

These data will be provided by DSSD and are expected to be available in June 2001.

3.	What are the housing unit coverage results for Puerto Rico from other housing unit coverage evaluation projects?

Methodology

We will look at other housing unit coverage evaluation projects for Puerto Rico results and include them in this report. A list of housing unit coverage evaluation projects that include Puerto Rico is included in the section of Related Studies and Operations (VIII).

V.	Limitations

This analysis does not take into consideration of matching errors and other non-sampling errors.

VI.	Cost Information

	FY '00	FY '01	FY '02
Number of FTEs for each Participating Division	DSSD 0.1	DSSD 0.5	DSSD 0.25
FTEs in Dollars (multiply total FTEs by $150,000)	DSSD $15,000	DSSD $75,000	DSSD $37,500
Contracts	none	none	none
Other	none	none	none

4

VII. Milestone Schedule

	Activity	Start Date	End Date
1.	Develop Study Plan	03/01/00	08/07/00
2.	Specify Data Needs	03/01/00	06/30/00
3.	Specify Other Activities	03/01/00	06/30/00
4.	Finalize Study Plan	08/07/00	09/30/00
5.	A.C.E. Housing Unit Matching and Follow-up	01/31/00	04/18/00
6.	A.C.E. Final Housing Unit Matching and Follow-up	03/07/01	05/22/01
7.	Delivery of the Data	06/01/01	06/30/01
8.	Start Analysis	06/01/01	06/30/01
9.	Start/End First Draft of the Report	09/01/01	12/31/01
10.	Roundtable Presentation	11/01/01	12/31/01
11.	Start/End Second Draft of the Report	01/01/02	06/30/02
12.	Prepare Final Report for Signature	07/01/02	12/31/02
13.	Report is Issued	12/31/02	12/31/02

VIII. Related Studies/Operations

Census 2000 Evaluation Project N.19. Housing Unit and Person Coverage Analysis.

Census 2000 Evaluation Project O.3. Coverage of Housing Units in the Census 2000.

Census 2000 Evaluation Project O. 7. Analysis of Person Coverage in Puerto Rico.

Census 2000 Evaluation Project O.9. Geocoding Error Analysis.

Census 2000 Evaluation Project O. 10. Housing Unit Duplication in the Census 2000.

Census 2000 Evaluation Project O.19. Analysis of Coverage of Housing Units int the Early Decennial Master Address File (DMAF) and Subsequent Census Coverage Improvement.

IX. References

Childers, Danny R., Accuracy and Coverage Evaluation: The Design Document, DSSD Census 2000 Procedures and Operations Memorandum Series, Chapter S-DT-1.

Childers, Danny R., Dual System Estimates of Housing Coverage, 1990 Decennial Census Preliminary Research and Evaluation Memorandum Number 230, 1993.

Urrutia, Maria E., Net Undercount Estimates from the Post Enumeration Survey (PES) of Puerto Rico, 1990 Puerto Rico PES Results Memorandum Number 1, 1993.